

## スペイン語作文添削における ChatGPT の活用可能性 ー日本の B1 学習者を対象とした事例研究ー

第 184 回関西スペイン語教授法ワークショップ (TADESKA) 例会

日時：2025 年 11 月 9 日 (日) 16:00 - 17:00

場所：Zoom を利用したオンライン開催

担当：平井素子・アイノア・ブハンダ

### El potencial del uso de ChatGPT en la corrección de redacciones en español: un estudio de caso con estudiantes japoneses de nivel B1CLXXXIV Reunión del

Taller de Didáctica de Español de Kansai (TADESKA)

Fecha y hora: Domingo, 9 de noviembre de 2025, de 16:00 a 17:00

Lugar: En línea (Zoom)

Moderadoras: Motoko HIRAI, Ainhua BUJANDA

\*\*\*\*\*

本ワークショップでは、自由テーマとして「日本のスペイン語学習者 (B1 レベル) による作文に見られる談話的規則性」に関する研究の中間報告を行いました。

授業内でセルバンテス文化センターのウェブサイトに掲載されている DELE B1 試験の Prueba 3 の課題に学生が取り組み、その 84 課題を ChatGPT-4 (有料版) がどのように評価・添削したかを分析し、その結果を共有しました。本研究にご協力くださった先生方に心より御礼申し上げます。

まず、ChatGPT に指示を与える際に使用した指示文の内容と、その結果 ChatGPT が作成した報告書を共有しました。続いて、ChatGPT による評価および添削の妥当性を精度 (Precision)、再現率 (Recall / Cobertura)、F1 スコアといった指標を用いて定量的に分析したところ、ChatGPT の評価は人間の評価よりも若干基準が緩く、本来であれば不合格となる作文も合格として扱っていることが明らかになりました。参加者の先生方からプロンプト設計に関して貴重な助言をいただきました。

次に、言語学的観点から ChatGPT による文法的誤りの指摘および解説の傾向を分析しました。綴りやアクセントの誤り、名詞の性・数および冠詞・形容詞との一致、動詞活用に関しては一貫して誤りを検出していることが確認されました (ただし、語頭の大文字・小文字の誤りは一部検出されませんでした)。また、gustar 型動詞や前置詞の誤用についても、ほぼ一貫して正しく指摘していました。

一方で判断に揺らぎが見られたのは、旅行中に食べた食事を「美味しかった」と述べる際の ser / estar の使い分け、そして 点過去と線過去 の使い分けでした。過去の出来事を現在形で述べている場合や、明らかに完結している出来事を線過去で述べている場合には誤

りとして検出されましたが、それ以外の文脈では誤りとして扱われませんでした。モデル文では正しい時制に修正されていましたが、ChatGPT の判断は文脈により揺らぎがあり、一貫性に欠ける点が見られました。また、稀に ChatGPT が誤って誤りを指摘し、誤った説明を与えるケースもありました。

この ChatGPT の報告書確認作業を通じて、日本人学習者に特有の文法誤りの偏りも明らかになりました。特に、旅行中に食べたものを述べる際に *ser* を用いる（正しくは *estar*）ケースが多かったこと、また「忘れられない思い出でした」「素晴らしい旅でした」といった旅行全体の評価文において、多くの学生が線過去を用いていたことが明らかになりました。これらの文は日本語では「描写」として表現されるため、学習者はスペイン語でも線過去を選びやすくなります。しかしスペイン語では、旅行という出来事全体を「完結した経験」として捉えるため、*Fue* のように点過去で表現するのが自然です。

この傾向は山村（2019）の指摘と一致しています。同氏は、スペイン語母語話者が点過去で述べる文脈において、日本語母語話者が評価を「出来事」ではなく「描写」として捉えるために線過去を過剰に用いてしまうことを、対照修辞学的観点（*retórica contrastiva*）から説明しています。

語用論的誤りや談話的規則性については現在分析を進めており、今後の学会等で結果を報告する予定です。

最後に、作文課題に AI（今回の場合 ChatGPT）を活用するかどうかについて参加者と議論しました。実際に分析した結果、生産性が大幅に向上するわけではないものの、フィードバックの質と添削の一貫性は明らかに向上することが確認されました。一方で、AI の進化は非常に速く、研究がそれに追いつかないという課題もあります。にもかかわらず、今回の議論を踏まえると、適切なプロンプトがあらかじめ整備されていれば、教育現場での活用も十分に検討し得るように思われました。

今回は作文添削における ChatGPT の使用について議論しましたが、練習問題の作成や授業設計など、他の分野にも応用可能です。今後は、教員の負担を増やすことなく学習効果を高める AI 活用法を共有していきたいと考えます。

最後に、当日使用したパワーポイント（PDF 版）とプロンプトを共有いたします。ご参加いただいた先生方、貴重なご意見をいただきありがとうございました。

En este taller, presentamos, por medio de una presentación en Power Point, un informe provisional de nuestra investigación acerca del tema de libre elección, «Regularidades discursivas observadas en composiciones de estudiantes japoneses de español (nivel B1)».

En él analizamos cómo ChatGPT-4 (versión de pago) ha evaluado y corregido 84 tareas de la Prueba 3 (publicada en la página web del Instituto Cervantes del examen DELE B1) realizadas por un grupo de alumnos en clase. Expresamos nuestro más sincero

agradecimiento a los profesores que amablemente han colaborado en esta investigación cuyos resultados se exponen seguidamente.

En primer lugar, suministramos unas instrucciones para guiar el ChatGPT y los informes resultantes generados por este. Posteriormente, analizamos la validez de la evaluación y corrección de ChatGPT utilizando métricas de Precisión, Cobertura y F1 Score. Estas revelaron que los criterios de evaluación de ChatGPT eran ligeramente menos estrictos que los de los evaluadores humanos, lo que daba lugar a que composiciones que normalmente se calificarían como suspensas, se consideraran aprobadas. A este respecto, recibimos valiosos consejos de los profesores participantes sobre el diseño de las indicaciones.

A continuación, analizamos los sesgos del ChatGPT a la hora de identificar y explicar los errores gramaticales desde una perspectiva lingüística. Se confirmó que ChatGPT detectaba de forma sistemática los errores ortográficos, de acentuación, de concordancia de género y número de los sustantivos, y de concordancia con los artículos y adjetivos, así como los errores de conjugación verbal (aunque no siempre detectaba los errores en las mayúsculas iniciales). Además, identificó casi siempre correctamente el uso agramatical de los verbos de afección interna y las preposiciones.

Sin embargo, fue ambiguo a la hora de distinguir entre *ser* y *estar* al describir las comidas consumidas durante un viaje como «deliciosas», y, al diferenciar entre el pretérito indefinido y el imperfecto. Si bien señaló correctamente los errores en el uso del imperfecto de indicativo para describir acontecimientos presentes o claramente concluidos, no identificó errores. Aunque las frases modelo se ajustaron al tiempo verbal correcto, los juicios de gramaticalidad del ChatGPT mostraron inconsistencia y falta de rigor cuando hubo de hacer intervenir al contexto. Además, hubo casos excepcionales en los que ChatGPT cometió errores y proporcionó explicaciones igualmente erróneas.

La revisión de los informes de ChatGPT también reveló sesgos en los errores gramaticales específicos de los aprendientes de ELE en Japón cuya primera lengua es el japonés. Concretamente, muchos estudiantes utilizaban *ser* en vez de *estar* al describir la comida que habían tomado durante el viaje. Asimismo, en la evaluación global de un viaje, como «Fue una experiencia inolvidable» o «Fue un viaje maravilloso», se constató un uso frecuente del imperfecto. Estas expresiones, que en japonés se formulan como descripciones, llevan a los aprendientes a seleccionar el imperfecto en español. Sin embargo, en español, el viaje se concibe como una “experiencia concluida”, por lo que debe expresarse en pretérito indefinido, con formas como Fue ....

Esta tendencia coincide con el análisis de Yamamura (2019). Desde la perspectiva de la retórica contrastiva, esta autora señala que los hablantes nativos de japonés tienden

a interpretar la evaluación de un viaje no como una “acción concluida”, sino como una “descripción”, lo que los conduce a un uso excesivo del imperfecto en este tipo de contextos.

Actualmente estamos llevando a cabo un análisis de los errores pragmáticos y las regularidades discursivas, cuyos resultados se presentarán en futuras conferencias académicas.

Por último, debatimos con los participantes la conveniencia de utilizar la IA (en este caso, ChatGPT) para la corrección de tareas de escritura. El análisis confirma que, aunque la productividad no aumentó significativamente, la calidad de los comentarios y la coherencia de las correcciones mejoraron claramente. Sin embargo, la IA evoluciona muy rápidamente, lo que plantea el reto de que la investigación tenga dificultades para seguirle el ritmo. No obstante, basándonos en este debate, parecía totalmente factible considerar su uso en entornos educativos, siempre que se preparen de antemano las indicaciones adecuadas.

En conclusión, ha sido objeto de discusión el uso de ChatGPT en la corrección de redacciones, aunque su aplicación se extienda a otras áreas, como la creación de ejercicios prácticos o la planificación de clases. De cara al futuro, nuestro objetivo es compartir métodos de utilización de la IA que mejoren los resultados del aprendizaje sin aumentar la carga de trabajo de los profesores.

Gracias a todos los profesores participantes por sus valiosos comentarios.